Multimodal Clinical Data Integration for Post-Discharge Risk Stratification: Ensemble Learning Approaches to Hospital Readmission Prediction

Dr. Beatriz Hernandez-Gomez, Professor of Industrial Engineering, Monterrey Institute of Technology and Higher Education (ITESM), Mexico

1. Introduction to Hospital Readmissions and the Need for Predictive Models

1. Introduction Hospital readmissions are a recurring issue that affects millions of people annually across the globe. When patients are sent home before necessary, are unable to afford prescribed drugs, forget to take them, are uninformed about symptoms that necessitate prompt medical attention, or are unable to contact their medical practitioner, the odds of them returning to the hospital increase. Readmitted patients have a variety of factors contributing to their readmission, including individual health, severe health-related determinants, environmental and behavioral characteristics, and genetics. The majority of strategies for preventing such occurrences recommend tailored interventions and regular follow-ups to high-risk patients after their discharge. The readmission burden not only affects patients but also increases healthcare system costs and medical expenditures. However, considering the time restraints, cognitive biases, and information overload in medical settings, it is problematic for caregivers to foresee potential post-discharge readmissions.

Various machine learning techniques can be employed to construct prediction models that authenticate and initialize the right clinical assistance to high-risk patients more consciously and accurately. Valuable information regarding patients' diseases, test results, and prescribed medications could allow healthcare professionals to diagnose and manage illnesses in an advanced manner. Various procedures such as logistic regression, decision trees, random forest, support vector machine, k-nearest neighbor, artificial neural network, and extreme gradient boosting can encode such details into a predictive model that helps to support in such cases. Nevertheless, building prediction models using the different methods of AI and ML is still considered a challenging task due to real-life care monitoring post-discharge, availability of incomplete datasets, imbalance ratio between readmitted and non-readmitted patient encounters, lack of organic integration of healthcare systems, data security and access issues, interpretability, and trust in machine learning models. Healing of such aftercare and monitoring prerequisites is possible with the help of advanced artificial intelligence techniques. AI techniques allow simulating human knowledge and reasoning to enhance several aspects of medical science and healthcare, which eventually leads to the rapid enrichment of patient care outcomes. Data processing in healthcare, incorporating this level of functionality, not only observes human-like capacities but also reflects thinking capabilities such as acquiring knowledge, understanding issue statements, discovering and comprehending complex ideas, inferring possibilities for investigating issues, and formulating ideas with the congregation of algorithms and complex data accurately with no risk proneness. AI techniques have steadily begun creating novel paradigms for understanding and ameliorating the systematic review challenges faced in hospital readmission and their statistical inference. These techniques have enabled digitalization in medical science to predict the occurrence of potential readmissions in various large-scale population studies at the post-discharge stage. This paper further provides potential AI-based medical science assisting healthcare administrators in minimizing their logistics costs and the burden caused to individuals by such aftercare services. Readmission analysis phenomena and considerations based on proposed frameworks are outlined across examples of diverse clinical conditions.

2. Foundations of Machine Learning in Healthcare

This paper centers around machine learning models developed to predict hospital readmissions in the form of non-technical, intermediate summaries. But first, let's lay some foundations.

The field of machine learning comprises the scientific study of computational statistical models, built either explicitly for predictive purposes or as a byproduct of data analysis, which learn to produce outputs or behaviors as a function of their inputs and internal parameters. Supervised learning, which encompasses the prediction or regression tasks, attempts to predict or model the probability distribution of some output, called the dependent or target variable, given some set of corresponding inputs, called the independent or predictor variables. Unsupervised learning aims to discover intrinsic patterns or structures in the data and does not involve the modeling of any specific output. Introduced to not only detect predictive signals that may be subconsciously recognized by experts but are unknown, unperceived, or imperceptible to human scrutiny, supervised machine learning has moved beyond probabilistic risk prediction to now interface directly with healthcare delivery, including automated decision-support frameworks.

In healthcare, many different types of algorithms are in use, many differ fundamentally, and some are explicitly designed and well-suited to learn and analyze complex and intricate data. Data quality is always paramount, but in the field of medical informatics, we must weigh these same standards against the pragmatic need for effective tools. For variables with essentially unknown biological relationships, more sophisticated and data-driven statistical methods are necessary, and hence these models tend to be more data-agnostic and better suited for use with disparate EHR data. Feature selection is important for model efficiency and interpretability because minimizing dimensionality often improves the performance of learning algorithms and reduces the time to train the model. Nevertheless, selecting interpretable features also informs the performance of more advanced and flexible models. Most models require feature scaling. Healthcare models face unique data and methodological challenges. Though many extra-EMRs offer enhanced and more detailed features, prospectively validating those models and workflows can be difficult. Furthermore, selecting the optimal model that generalizes well across institutions is difficult.

This crossroads between healthcare and technology is the site of many exciting advancements that have the potential, if leveraged properly, to dramatically enhance patient management. Therefore, knowledge in this area is valuable to - and may benefit - all stakeholders involved. Later sections will detail more advanced models for predicting hospital readmissions.

3. Data Sources and Preprocessing for Hospital Readmission Prediction

Predicting hospital readmission events from patient data is an interesting challenge, and there are a number of datasets that can be used to train models. The most common type of data is unstructured text describing clinical records, which is typically the free text found in discharge notes and electronic health records. Other sources of useful data include demographic information and survey/quality of life tools filled in by the patient to assess patient mood and health-related quality of life. The first step in a machine learning pipeline after deciding what data to collect and what variables to store in the data is to preprocess the data. Preprocessing includes ensuring data is clean and normalized appropriately; this is important as incorrect preprocessing can corrupt results. Due to the variability and incomplete nature of patient information, a number of techniques are used to ensure that the results are not skewed by patient records that are not fully filled in. One of the common methods is to only include fully populated records; however, it has been shown that using advanced techniques for handling missing records can actually improve the performance of the classification model considerably. Research standards in this area often involve removing data points with missing measurements or using the first observation carried backwards as an imputation for the missing value. By ensuring a good data pipeline, we can concentrate on the main task of the paper, which is how the performance of the model can be improved by integrating different types of data. As well as the advantages in prediction performance, well-prepared data creates an avenue for ethical AI. Briefly, ethical AI represents the extensive pipeline activity that occurs to protect personal health information from misuse and unintended consequences. There are many sources of data that can be used in a hospital readmission model. The largest source of information is the clinical records. These records are kept in an electronic format and are often referred to as electronic health records. The data provide an undisputable record of the information that was provided to a patient and the treatments the patient procured. It also provides an outlet for communication between healthcare professionals that might take over the care of a patient. The records are collected as part of regular practice and normally serve to describe a single consultation or inpatient stay.

4. Key Machine Learning Models for Predicting Hospital Readmissions

We discussed three primary machine learning models for the prediction of hospital readmissions. Logistic regression is used for binary outcome prediction and for the identification of important biomarkers associated with hospital readmissions. Random forest is utilized for assigning a predicted probability score, identifying important predictors, and handling correlated features. It is super fast to train a model that can be used to serve as an initial solution before creating a number of boosting models with various features as input. Gradient boosting machines typically perform better and have higher accuracy than random forest at the cost of longer training times. They also provide inbuilt feature selection.

Machine learning models are trained according to the recent past medical history of diabetic patients to predict if they will be readmitted within a short time period after being discharged from the hospital. Logistic regression, random forests, and gradient boosting machines are evaluated based on various criteria such as the area under the Receiver Operating Characteristic curve and other parameters such as accuracy, precision, and F1 score. We also evaluate the models based on their interpretability and the computational effort needed to set up and optimize these models by fine-tuning their hyperparameters. The benefits for predictive performance and for clinical management that derive from the capacity of an algorithm to identify patients at high risk of unplanned readmissions can be realized only if the model is sufficiently reliable, scalable, and interpretable. Therefore, in deciding which model to use in which scenario, one should consider their predictive ability and other criteria that are specifically tied to their ease of use and usability in clinical settings. In this sense, interpretability has been defined as the capacity of a model to create an understanding in an interpretable or reasonable manner.

4.1. Logistic Regression

Logistic regression, one of the simplest and most widespread techniques for predicting hospital readmissions, turns out to be a natural choice in many healthcare machine learning applications. For outcomes that are binary in nature, such as readmitted and not readmitted patients, logistic regression forms an appealing choice. The mathematical framework of logistic regression consists of a sigmoid function that maps the predicted value to [0, 1], converting it into a probability. With features as input, this converted value represents the probability of occurrence of the event (in this case, the likelihood of readmission). Furthermore, the model and all related iterations can also predict the odds ratio from this logistic curve. Another advantage of logistic regression is that the model yields AUC/ROC values akin to tree models, making it a model with efficient interpretability and performance.

Logistic regression has been extensively used alongside various problems in predicting health-related outcomes. A logistic regression model predicted 30-day unplanned emergency department visits with accuracy and AUC of 73.3% and 0.74. In another attempt, the effort of predicting readmissions in chronic heart failure (CHF) patients was undertaken using logistic regression. Various accuracy scores for the predictive model were attained, with the best one demonstrating an out-of-sample accuracy of 60.5%. Others developed a readmission model for heart failure (HF) and acute myocardial infarctions (AMI) (with predictions at multiple intervals), chronic obstructive pulmonary disease (COPD), and pneumonia. Establishing a linear baseline model for comparison, logistic regression model accuracy was evaluated after cross-validation (where applicable), showing scores between approximately 58% and 60%. The logistic regression model was employed to predict a variety of healthcare outcomes; however, it is essential to note that this model, as a linear classifier, has limitations in its ability to handle high-dimensional and complex settings.

When choosing to use logistic regression, it is important to recognize the strengths and limitations of the model. First, logistic regression is a very efficient algorithm and is often sufficient for generating accurate predictions. Logistic regression is also a very interpretable model, which makes it a good candidate for problems for which interpretability is necessary or desirable. Finally, because logistic regression outputs probabilities by default, it can be directly used to assign patients to high- and low-risk groups. Although these strengths make logistic regression a good fit for many problems, it is not perfect for all applications. Logistic regression works best with a moderate number of features that themselves have a linear relationship with the predicted event. Moreover, logistic regression will struggle with sparse data if the positive class has very few samples. At its core, logistic regression models the relationship between the dependent variable and one or more independent variables by estimating probabilities based on data from a previous sample. The aim is to create a predictive model where the predicted output is the probability of the target variable belonging to the default class, termed the success class. This makes logistic regression an approach appropriate for binary classification problems. Overall, logistic regression is a useful tool in your prediction toolbox to use as part of a larger ensemble of predictive models.

4.2. Random Forest

Random forests consist of an ensemble of decision trees and are a popular choice for predictive modeling due to their robust performance. Decision trees work by recursively performing binary splitting on the input space, creating a flowchart of binary decisions that lead to a precise outcome. The primary advantage of random forests is that they are a vote from a multitude of weak learners, a collection of decision trees. This can improve predictive accuracy and reduce overfitting associated with decision trees in comparison to an individual strong learner.

Random forests can model any nonlinear relationship or interactions between features; it should naturally be a useful algorithm when modeling healthcare data. Various studies have demonstrated the effectiveness of random forests in predicting hospital readmissions. A random forest model reported an AUC value of 0.732 on the test set in depression patient data of 12,497 patients. The random forests algorithm is the most popular and common method for predicting hospital readmissions. This considerable interest and focus in the healthcare space is due to the algorithm's robust performance and ability to provide highly interpretable feature importance inclusions. This model is preferred among clinicians and data scientists, not only for its performance but primarily for the interpretability of feature importance that can contribute to improving post-discharge care.

The primary disadvantage of this algorithm is that in large datasets it can be computationally expensive, especially when hyperparameters like 'number of trees' and 'max tree depth' are not optimized. Additionally, the random forests algorithm also struggles with model interpretability. The hyperparameters commonly tuned with random forest models include the number of estimators, max tree depth, and minimum leaf samples. Evaluation metrics common for random forests include AUC-ROC, AUC-PR, F1-score, and precision. F1-score is a particularly important evaluation metric for this project in comparison to accuracy because it is also important to identify as many positive cases of POC mortality as possible; therefore, the recall needs to be quite high.

4.3. Gradient Boosting Machines

Also known as GBM, Gradient Boosting Machines are yet another powerful class of techniques, often used for predicting patient readmissions. The basic idea behind the GBM algorithm is that it builds weak learners sequentially, whereby each learner concentrates on the remaining errors. It seeks to minimize the error by grappling with the weak points of the previous learners. The final prediction, which is conducted through additive modeling, is referred to as an ensemble. GBM is hence a prominent ensemble technique that seeks to offer high accuracy for predictions. State-of-the-art GBM uses state-of-the-art weak learners for better predictive performance.

There are several parameters that can be optimized in GBM models for inpatient readmission forecasting. These include: (1) the learning rate, which boosts the contribution of each weak learner; (2) the number of learners or estimators; (3) the depth of the tree; and (4) the number of randomly selected variables from the data for splitting in each node. All of the aforementioned parameters can be optimized to receive the best predictive results from the GBM algorithm, and their impact should be evaluated by adjusting their values across a grid. For demonstration, GBM was utilized to perform predictions for diabetes and heart failure, resulting in improvements in AUC values by 1.50% and 1.00% respectively, as compared with regression model baseline predictions. One limitation of GBM is that it is sensitive to noisy data, and when not properly tuned, it may result in the overfitting of the training data. In the healthcare domain, GBM is often used as an ensemble technique alongside Random Forest, with empirical success. Understanding GBM will allow for comparing and potentially improving the performance of multiple models in ensemble.

5. Evaluation Metrics for Model Performance

This section explains how we can evaluate the performance of machine learning models. Let the confusion matrix of a pre-trained predictive model compute the true negative (TN), false positive (FP), false negative (FN), and true positive (TP) predictions. Then, we can use these four values to calculate the following evaluation metrics: accuracy (ACC), precision (PREC), recall (RE), F1-score (F1), Area Under the Curve (AUC), and Receiver Operating Characteristic (ROC) to analyze the performance of a pre-trained predictive model.

These metrics have been extensively used for training models that predict hospital readmissions. We briefly introduce them here and describe the importance of using them in practice. The confusion matrix is how we conclude that a classifier's outcome is a recommended tool, allowing us to interpret the model's efficacy. If our evaluation is based on ACC and PREC, we may commit to the superiority of the LR model due to its prevalence of TN. On the other hand, FN were oversampled in the ML model, skewing its RE, even though it has a prevalent TP and fewer FP counts. An evaluation based on RE shows the contrary.

It is important to note that the use of one of these three metrics can be subject to bias towards a specific model. Thus, to avoid biased interpretation, it is important to have a balance of the three measures when assessing model performance. One could mistake the models by their accuracy alone, which is not suitable for clinical diagnosis. ACC has a strong propensity towards the majority class and represents classifiers that underestimate the minor class. It is generally used to identify diagnostic routines. However, one cannot take for granted the RE and PREC only as it may lead to higher sensitivity at the expense of a lower FP rate and vice versa. The terms false alarms and missed detections can be extensive as per their clinical meanings, such as medication or preventive decisions.

6. Real-World Applications and Case Studies

In this section, we will delve deeper into the real-world applications and case studies of machine learning models for predicting hospital readmissions. Many papers do not only focus on methodological contributions but also on implementation and improvements in specific healthcare settings. Researchers observe the potential of predictive modeling techniques based on big data analytics applied to discharge planning, which makes it possible to discriminate and identify high-risk patients for suitable treatment. This approach would lead to the reallocation of healthcare resources, early detection of adverse events, extensive monitoring of patients' health states, decision-making, performance analysis, and overall quality improvement. In terms of reducing readmission, there are hospitals reporting the ability to cut down the first 30-day all-cause readmission rate by applying predictive readmission algorithms. The models are coupled with or available via electronic health records systems.

Some applications of machine learning and analytical algorithms that use structured and unstructured data to predict hospital readmissions after the discharge of sicker adults are analyzed. The projects aim to operationalize the predictive models to support decision-making among healthcare providers. For example, healthcare assignments can be organized so that nurses and social workers systematically prioritize visits and telephone calls to patients at higher risk, support new modes of surveillance, intervention, communication, and resource allocation, and, in doing so, reduce disciplines that no longer benefit from scarce resources to complete this task. Hospital units can also use these approaches to coordinate and manage bed availability, discharge planning, and patient flow. The described operational applications pertain to geriatrics but are relevant to general adult medical and surgical care. Lessons learned from these predictive projects are the importance of classifying the objectives of these algorithms as a ranking/sorting function, feedback loops, and the importance of monitoring effects.

7. Ethical Considerations in Using AI for Patient Care

The development of AI-based models for improving patient outcomes and healthcare delivery brings ethical considerations parallel to improvements in patient care. Pertinent questions include patient monitoring and protection, model transparency, and distribution. General ethical philosophies can be helpful guides in these circumstances. Importantly, a program founded upon machine learning analyzes data from patient charts, making privacy and safety major concerns. As more sophisticated clinical models are developed, the potential benefits must be weighed against potential risks such as biased algorithms. There is a need for fair and transparent testing of new AI models to ascertain their utility in clinical care, rather than an unchecked proliferation of new AI models rushed to implementation with promises of improving care and hacking the healthcare system.

Regarding privacy and data security, patients should always be made aware when their data are being transferred, and permission for the funds and federating hospitals to use de-identified healthcare records is paramount. There should also be data encryption to ensure proper safety measures. With a reliance on electronic medical records and AI-enabled healthcare systems, such as in readmission prediction, such data transfers are often necessary, as is continued research into developing secure and compliant electronic healthcare systems. Moreover, there should be no legal consequences for the patient in the event of a breach of their de-identified healthcare record in the hands of the AI-intervening hospital system. In light of the ethical concern, stakeholders are called to explain to institutional review boards, patient representatives, and clinical end-users why morally problematic models were abandoned, and the timeline of such revelations. In order to address this ethical concern, our AI for predictive modeling is used directly at the institutional level. It is a feedback tool used to enlighten the setting of care at an intensive hospital for a few days. As our study includes only de-identified patient-facing data, it does not require consents, notwithstanding mandatory research reporting.

8. Conclusion

In this paper, we have discussed machine learning models for predicting hospital readmissions. Hospital readmissions have received increasing attention in the research community, governments, and healthcare organizations around the world. The availability of affordable, powerful, and advanced tools opens up a new door for the development of more accurate models designed to predict patients at high risk of readmissions, even for a longer time span. Developing and deploying models for predicting hospital readmission should address sensitive issues, including the issue of model evaluation and validation; the fairness and bias of machine learning models; and issues of data privacy and security. Furthermore, the results in this paper are not intended to replace the treatment by clinical care providers. There are numerous things to do, and the ultimate impact of machine learning is a multimodal approach that involves multiple stakeholders in a care team. Regardless of the further enhancements and research, we note the possibilities of machine learning in predicting hospital readmission and care for all patients, especially those at greatest risk. In conclusion, in this paper, we have explored machine learning models for predicting hospital readmissions. Machine learning models have various applications for improving healthcare. Such models can be used to increase flexibility and accuracy when searching for patients who have the highest risk of being readmitted because internal immune prediction techniques can capture complex patient problems at admission that can benefit from post-discharge medications, support services, or other services. It is becoming increasingly apparent that readmission models can contribute to efforts to achieve the goal of improving patient outcomes in hospitals and nursing homes, representing one of the major challenges for medical science, data science, and clinical practice. The possibility of employing machine learning models to improve patient outcomes necessitates ongoing analysis and investment by practitioners and researchers from various disciplines.